HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

The Links Between Analytical Thinking Skills, Decision-Making Characteristics and Temperament

University of Helsinki
Faculty of Educational Sciences
General and Adult Education
Master's Thesis
April 2019
Sanni Bachmann

Supervisors: Heta Tuominen
Markku Niemivirta

HELSINGIN YLIOPISTO

Tiedekunta Kasvatustieteellinen tiedekunta		
Tekijä Sanni Bachmann		
Työn nimi Analyyttisen ajattelutaidon, päätöksenteko-ominaisuuksien ja temperamentin väliset yhteydet		
Oppiaine Yleinen- ja aikuiskasvatustiede		
Työn laji/ Ohjaajat Pro gradu -tutkielma / Heta Tuominen, Markku Niemivirta	Aika Huhtikuu 2019	Sivumäärä 40 s + 3 liites.

Tiivistelmä

Tutkimuksen tavoitteena oli tarkastella analyyttisen ajattelutaidon, valikoitujen päätöksenteko-ominaisuuksien (aikapreferenssi, riskipreferenssi) sekä temperamentin välisiä yhteyksiä. Laajasti sovellettu analyyttisen ajattelutaidon testaamisen väline on Cognitive Reflection Test (CRT; Frederick, 2015). Testi mittaa rationaalista ajattelutaitoa, joka on erotettavissa älykkyydestä sekä muista kognitiivisista kyvyistä (Toplak, West, & Stanovich, 2011). Mikäli vahva korrelaatio analyyttisen ajattelutaidon ja eräiden muiden ominaisuuksien välillä kyetään todentamaan, auttaa tämä paitsi tunnistamaan yksilöt, joilla hyvä analyyttinen päätöksentekokyky on, mutta myös selventämään, mistä analyyttisessä ajattelutaidossa on ilmiönä kyse.

Aiemmissa tutkimuksissa on havaittu yhteyksiä yksilön analyyttisen ajattelutaidon (mitattuna CRT-testillä), sekä aikapreferenssin ja riskipreferenssin välillä yhdysvaltalaisilla yliopisto-opiskelijoilla. Tässä tutkimuksessa aikapreferenssiä mitattiin Consideration of Future Consequences -testillä (CFC-14; Joireman, Shaffer, Balliet, & Strathman, 2012). Suomalaisilta yliopisto-opiskelijoilta kerätyllä aineistolla pyrittiin replikoimaan aiempia tuloksia (Frederick, 2015) riskipreferenssin osalta. Tutkimus myös selvitti niin analyyttisen ajattelutaidon kuin aikapreferenssinkin mahdollisia yhteyksiä neurobiologiseen temperamenttiin (Rawlings, Tapola, & Niemivirta, 2017). Jälkimmäisiä yhteyksiä ei ole tiettävästi aiemmin tutkittu.

Tutkimusaineisto kerättiin Maanpuolustuskorkeakoulun 135 ensimmäisen vuoden opiskelijalta. Opiskelijat vastasivat kyselylomakkeeseen internet-linkin kautta. Mallien rakenteellista validiteettia tutkittiin eksploratiivisen faktorianalyysin avulla. Mittareiden välisiä korrelaatioita tarkasteltiin, ja regressioanalyysin avulla analysoitiin, miten temperamentti ennustaa aikapreferenssiä.

Odotusten vastaisesti yhteyksiä analyyttisen ajattelutaidon ja aikapreferenssin tai riskipreferenssin välillä ei löytynyt. Aikapreferenssin ja temperamentin välillä havaittiin kuitenkin kiinnostavia yhteyksiä. Tulevaisuudessa olisikin mielenkiintoista tutkia lisää neurobiologisen temperamentin mahdollisia vaikutuksia siihen, kuinka suhtaudumme kauempana tulevaisuudessa odottaviin ja välittömiin seurauksiin, toisin sanoen, kuinka kaukonäköisiä olemme.

Avainsanat Analyyttinen ajattelu, päätöksenteko, aikapreferenssi, riskipreferenssi, temperamentti	
Säilytyspaikka Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet)	

HELSINGIN YLIOPISTO

Faculty Educational Sciences		
Author Sanni Bachmann		
Title The Links Between Analytical Thinking Skills, Decision-Making Characteristics and Temperament		
Subject General and Adult Education		
Level/Instructors Master's Thesis / Heta Tuominen, Markku Niemivirta	Month and year April 2019	Number of pages 40 pp. + 3 appendices

Abstract

The purpose of this study was to look for relationships between analytical thinking skills, two selected decision-making characteristics (time preference, risk preference) and temperament. A commonly applied test to detect individual differences in analytical thinking is the Cognitive Reflection Test (CRT; Frederick, 2005), which captures important characteristics of rational thinking, different from intelligence and other cognitive ability (Toplak, West, & Stanovich, 2011). If a strong correlation between the ability to analytical thinking and some other types of characteristics was found, it would help to identify individuals with good analytical decision-making skills and to clarify the phenomenon of analytical thinking.

This work continued the exploration of relations between individual differences in analytical thinking, assessed as performance in CRT, and selected psychological characteristics: time preference and risk preference, where a link has been detected in previous research on university students in the U.S. In this study, time preference was assessed as performance in the Consideration of Future Consequences Test (CFC-14; Joireman, Shaffer, Balliet, & Strathman, 2012). With data gathered from Finnish university students, this study attempted to replicate the findings of the original study (Frederick, 2005), where a relationship between the CRT and risk preference was detected. The current study extended previous research by examining also the relation of analytical thinking skills and neurobiological temperament (Rawlings, Tapola & Niemivirta, 2017). The relation of time preference and temperament was also studied.

The participants in the present study were 135 first-year students from the National Defence University in Finland. The students completed a self-report questionnaire via a weblink. An exploratory factor analysis was performed to test the construct validity of the models applied. The correlations between the tests were examined and by using standard regression analyses the effects of the temperament on the time preference were analysed.

Contrary to expectations, no relation between the CRT score and time preference, risk preference or temperament were detected. However, interesting links between time preference and temperament were found. Therefore, future research could investigate further how neurobiological temperament sets our concerns with future and immediate consequences, that is, our farsightedness or the lack of it.

Keywords

Analytical thinking, decision-making, time preference, risk preference, temperament

Where deposited

Helsinki University Library – Helda / E-thesis (theses)

Contents

List of figures and tables

List of appendices

Abbreviations

CFC	Consideration of Future Consequences
CRT	Cognitive Reflection Test
SP	Sensitivity to punishment
SRe	Interindividual reward sensitivity
SRiNS	Intraindividual reward sensitivity – novelty-seeking
SRiPE	Intraindividual reward sensitivity – positive expressiveness

1. Introduction

The aim of this study is to examine relationships between skills in analytical (rational) thinking and selected psychological characteristics, that were assumed to correlate with it. Assessment of decision-making skills, such as analytical thinking, can be important for example, when executives and directors are recruited – or students for a special field of study are selected. Attempts to develop such tests have been made. However, creating these tools is a challenge: as soon as the individual tries to solve the test problem or is given the correct answer, the problem and answer risk to be memorized. Tests often attract interest from the public, and the test problems become well-known, which makes them not effective for continuous use (Toplak, West, & Stanovich, 2014). This challenge has attracted interest to look for possible relations between analytical thinking skills and other psychological characteristics. If a strong correlation between the ability to analytical thinking and some other types of characteristics was found, it would be easier to identify individuals with good analytical decision-making skills.

What is analytical thinking after all? Rationality, which depicts the ideal model of behaviour, is not easy to assess as a subject of study. However, it has long been studied in different fields, including philosophy, cognitive science, psychology, economics and judgement and decision-making (JDM) (Stanovich, 2009). In JDM the classic standard for rationality is utility theory, a cornerstone of modern economics. Utility theory sees decision-making as a rational process, purely based on calculating the highest expected utility. However, as empirical studies have shown, utility theory does not successfully model the human decision-making behaviour in real-life circumstances, not even in the simplest situations (Kahneman & Tversky, 1979; Tversky & Kahneman, 1974). In response to this, Amos Tversky and Daniel Kahneman developed the Nobel prize winning prospect theory to explain why so often people are risk averse in the domain of gains, but risk-seeking in the domain of losses.

Why did standard economic models of human decision-making fail to explain human behaviour? One of the reasons is their rejection of emotions. Even though the study of emotions within JDM is relatively young, modern science recognises the importance of the emotional system in decision-making (Clore & Huntsinger, 2007; LeDoux, 1996; Lerner, Li, Piercarlo, & Kassam, 2015; Loewenstein & Lerner, 2003). It is from the bases of our emotions, that we make the choice to avoid or approach. If our emotional system is

not working properly, we fail to make good decisions (Bechara, 2004; Damasio, 1994). Due to our emotions, we also tend to reject unfair offers, even when this is not rational behaviour from the perspective of utility theory (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Although emotions are necessary for and inseparable from decision-making, sometimes emotions can be misleading. Even though the reliance on heuristics ("gut feeling") might have been an adequate strategy for our hunter-gatherer ancestors, and intuitive judgment is still useful in everyday life, many decision-making situations of the modern world require a different type of approach.

In cognitive psychology, cognitive neuroscience and other fields studying thinking, two types of cognitive processes are often distinguished: thinking can be divided into intuitive (heuristic) and analytical (rational). Intuitive thinking is liable to cognitive biases which can lead to severe errors of judgment and cause profound consequences. Classic examples of cognitive biases, mental processes of simplifying complex data, are for example anchoring and availability (Kahneman, 2012). Anchoring describes our tendency to rely on the first piece of information offered (the "anchor") when making decisions. Also, we tend to take into consideration only the information available, ignoring that information is missing. Are some individuals more analytical in their decision-making than others, and some more prone to rely on heuristics? Several studies have detected individual differences in these two thinking styles or systems. It is not known, to what extent these tendencies are innate, but they are presumed to correlate with other individual differences, such as cognitive styles.

The purpose of this study is to analyse the relationship between analytical thinking skills and two decision-making characteristics: time preference and risk preference. In this study, analytical thinking skills are assessed as performance in the Cognitive Reflection Test, (CRT; Frederick, 2005), based on dual-process theory described above. An individual's attitude toward future outcomes has long been linked to rationality and decision-making. This consideration of temporal consequences has been measured with a variety of psychometric methods under the title of time preference or temporal discounting. A relation between CRT and time preference has been detected in previous studies, as well as a link between CRT score and risk preference, that is, individual differences in risk attitude (Frederick, 2005; Toplak et al., 2014). Decision-making is often about deciding whether to approach or avoid, a tendency which can vary according to temperament. For that reason, relationship between the ability to analytical thinking and temperament, measured as punishment and reward sensitivities (Rawlings, Tapola & Niemivirta, 2017), is also studied. The original studies of analytical decision-making skills were conducted

on students from prestigious universities such as Harvard, MIT, and Princeton. This study is replicated in the higher education context, the participants being students of the National Defence University in Finland, where students are selected after a multi-tier selection process.

2. Analytical and intuitive thinking

2.1. Dual-process theories

The idea of the mind being composed of different subsystems is nothing new – already Plato divided the soul into three agents: reason, spirit, and appetite (Evans & Frankish, 2009; Stanovich, 2011). The view of a fundamental duality in the human mind developed into modern dual-process theories after the cognitive revolution in psychology, during the 1960s and 1970s (Evans & Frankish, 2009). The modern dual-process theories divide cognitive processes into intuitive (heuristic) and analytical (rational). Various alternative views and terms for the dichotomy have been proposed, for example, Mischel distinguishes between hot and cool systems (Metcalfe & Mischel, 1999; Stanovich & West, 2000). Stanovich (1999) introduced the labels System 1 and System 2, which have since then become popular. Later Stanovich suggested that Type 1 and Type 2 terminology should be applied, because the term 'system' refers to a singular system, while processing might result from operations of a variety of subsystems (Stanovich, 2011). Evolutionary theorizing, as well as neurophysiological research, have given support for the dual-system models. System 1 is assumed to be relatively old from the evolutionary perspective and shared with other animals, whereas System 2 is considered as more recently evolved and typical to humans (Gilhooly & Murphy, 2005).

According to the dualistic theory, people are prone to save energy, making most of our thinking automatic (heuristic). We often behave like cognitive misers: we choose the quick response or solution that comes to mind instead of suppressing the erroneous response and thinking further (Stanovich, 2009; Toplak et al., 2011; Tversky & Kahneman, 1974). This Type 1 thinking is unconscious and fast, whereas the deliberative Type 2 processes are conscious and slow (Evans & Frankish, 2009). Automatic thinking is well needed and accurate most of the time (Evans & Stanovich, 2013; Kahneman, 2012). However, the intuitive mind has the tendency to commit thinking errors (cognitive biases; Stanovich 2009; Tversky & Kahneman, 1974). Therefore, when making decisions of un-

usual importance, ideas produced by System 1 should be analysed by System 2. Nevertheless, contrary to a common misconception, lack of emotions is not an attribute of System 2 (Kahneman, 2012). It is nowadays well understood that emotional processes are involved in all decision-making (Bechara, 2004; Lerner et al., 2015).

Systematic differences in the tendency to biased thinking exist between individuals (Stanovich, 2009). It is not known to what extent the differences are learned or innate. Research has shown that intelligence is only a moderate predictor of rational thinking and that some rational thinking skills can be dissociated from intelligence (Stanovich, 2009; Stanovich 2011). Rationality is operationalized extensively in modern cognitive science, but none of these operational measures are assessed on common IQ tests (Stanovich, 2009). Intentions to assess a rationality quotient (RQ) test have been made (Stanovich, West & Toplak, 2016).

2.2. Cognitive Reflection Test

To test individual differences in analytical thinking skills, Frederick introduced the Cognitive Reflection Test (CRT; Frederick, 2005). The original CRT consists of three problems that first generate an intuitive response, which is incorrect. For example: "*A bat and ball cost $1.10. The bat costs one dollar more than the ball. How much does the ball cost?"* This problem evokes the intuitive answer of 10 cents, the correct answer being 5 cents. The CRT measures whether an individual is satisfied with the intuitive response, or if he or she will keep searching for a better solution and is then capable of producing the correct answer. Frederick analysed the writings in the margins of the answer forms of his studies and noticed that the incorrect, intuitive answer was often considered, even though it was not chosen as the final solution. When asked about the difficulty of the test problems, the respondents not scoring high found them easier than the respondents who were capable of solving them. The CRT problems differ from classic insight problems, which usually do not trigger an intuitive, incorrect response (Gilhooly & Murphy, 2005; Toplak et al., 2011).

The CRT was administered to 3,428 respondents in 35 separate studies (Frederick, 2005). Many of the respondents also took other measures of cognitive ability, such as the Wonderlic Personnel Test (WPT) and the Need for Cognition scale (NFC), and self-reported their American College Test (ACT) and Scholastic Achievement Test (SAT) scores, as well as tests of time preference and risk preference (see 3.1. and 3.2.). The

relations of the CRT with these other tests were then examined and showed positive and significant correlations. Between CRT score and SAT scores, Frederick (2005) detected a correlation of 0.44, and Obrecht, Chapman and Gelman (2009) a correlation of 0.45. When the various cognitive measures were correlated with composite indices of decision-making characteristics formed from the time preference and risk preference items, the results showed that the CRT was either the best or the second-best predictor of intertemporal choice and choice under uncertainty.

Several studies have examined the relation of the CRT and other measures of cognitive ability. The findings of different studies show a significant correlation between the performance on the CRT and several cognitive biases, non-superstitious thinking and general numeracy (Fernbach, Sloman, Louis, & Shube, 2013; Liberali, Reyna, Furlan, Stein & Pardo, 2012; Mata, Ferreira & Sherman, 2013; Moritz, Hill, & Donohue, 2013; Oechssler, Roide & Schmitz, 2009; Pennycook, Cheyne, Seli, Koehler & Fugelsang, 2012; Shenhav, Rand, & Greene, 2012). In contrast, no relation between CRT performance and the degree of encounter bias was found (Obrecht, Chapman & Gelman, 2009), and only little relation between CRT results and the choice of a high-expected-value gamble was detected (Campitelli & Labollita, 2010).

Toplak, West and Stanovich (2011) demonstrated that the CRT was a better predictor of rational thinking than the several measures of intelligence and executive functioning (inhibition, updating and set shifting; Stanovich, 2011) used in their study. They formed a composite variable of 15 separate rational thinking tasks including three executive-functioning tasks, a cognitive ability test (Wechsler Abbreviated Scale of Intelligence, WASI), heuristics-and-biases tasks, two sets of syllogistic reasoning tasks and thinking dispositions. Thinking dispositions or cognitive styles are measures of the functioning of the reflective mind – concerning beliefs, belief structures, attitudes toward forming and changing beliefs, and goals (Stanovich, 2011). Thinking dispositions measure individual differences in goal management, epistemic values, and epistemic self-regulation: examples of thinking dispositions include the tendency to seek various points of view before coming to a conclusion, the disposition to think extensively about a problem before responding or the tendency to think about future consequences before taking action. Time preference (see 3.1.), a characteristic studied in this thesis, is often considered as a thinking disposition (Toplak et al., 2011). The CRT, consisting of only three items, was found to be a better predictor of rational thinking than either measures of intelligence or measures of executive functioning (Toplak et al., 2011, 2014).

Only a few studies have attempted to examine the relation between performance in the CRT and performance in “real-life decisions”. A correlation between CRT scores and success in a real-life simulating task was found when managing inventory with a computer program (Narayanan & Moritz, 2015; Moritz, Hill & Donohue, 2013) and in a computer-based banking game (Kiss, Rodriguez-Lara, & Rosa-García, 2016).

The original CRT is a three-item test. All three CRT questions (CRT1-CRT3) have appeared frequently in literature and social media, which induces a risk of familiarity. Therefore, a similar seven-item test, introducing four supplementary questions, was designed by Toplak et al. (2014). For the revised version, problems CRT4 and CRT5 were supplied by Frederick. CRT6 was adapted from Dominowski (1994), and CRT7 was created by Toplak and her colleagues.

3. Intertemporal choice and choice under uncertainty

3.1. Time preference

The influence of dual-system models extends to cognitive neuroscience, trying to understand the psychological processes of self-control, modulating impulsivity and far-sighted behaviour (McClure & Bickel, 2014). A long-standing hypothesis states, that there is a link between deliberative control of behaviour and the executive system in the brain, as well as between the automatic mode of behavioural control and brain reward areas. Self-control (delay of gratification, effortful control, willpower, executive control, time preference, self-discipline, self-regulation, ego strength) is the idea of effortful regulation of the self by the self (Duckworth, 2011). Self-control is related to conscientiousness, a personality trait that includes responsibility, industriousness, and orderliness (Moffitt et al, 2011). Behavioural studies have revealed, that self-control skills, ability to inhibit impulsions, is related to almost all “success in life” (Benjamin, Brown, & Shapiro, 2013; Moffitt et al., 2011; Schlam, Wilson, Shoda, Mischel, & Ayduk, 2013). Ability to self-control predicts for example savings behaviour, income, physical and mental health and lack of criminal convictions (Duckworth, 2011).

Self-control, the ability to resist the temptation to immediate reward to receive a larger later reward, is often examined under the title of time preference (temporal discounting, time discounting, delay discounting). Presumably, the most widely known tests for self-

control were conducted by Mischel in the 1960s and 1970s, often referred to as the "marshmallow tests" (Mischel, 1974). Since then countless studies have found a relation between cognitive ability and patience (Benjamin et al., 2013; Parker & Fischhoff, 2005; Shoda et al., 1990), and delay discounting has become a standard measure of self-control. When adults are being tested, no tempting marshmallows are offered for the participants, but time preference is measured with psychometric methods (Mazur, 1987; McClure & Bickel, 2014; Rachlin, Raineri, & Cross, 1991). Respondents are asked to compare rewards of different amounts available at different time delays, and then choose their preferred option. For example: "*Would you prefer 3400 euros this month or 3800 euros next month?"*; or "*Would you prefer 50 euros immediately or 1000 euros in 10 years?".* Neuroimaging has supported the findings of automatic and deliberative control processes produced by behavioural studies, yet it is not completely known how these processes relate to neural systems (Evans & Stanovich, 2013; McClure & Bickel, 2014).

It is interesting to examine the relation of rational thinking skills and the ability to control impulsivity. The three CRT problems are "easy" in the sense that their solution is easily understood when explained (Frederick, 2005). Therefore, it is expected that the difficultness lays in the respondent's capability to suppress the intuitive, incorrect answer that comes into mind impulsively. One of the CRT problems, the bat-and-ball puzzle, was adopted from an experiment on cheating (Nagin & Pogarsky, 2003), where the results revealed that the respondents giving the intuitive answer were less patient than respondents producing the correct answer.

Frederick (2005) did examine the relation between CRT score and time preference, measuring whether an individual prefers an immediate reward if proposed with two alternatives. The choice between an immediate reward and a larger delayed reward was significantly associated with the test score. For instance, the respondents scoring low in the CRT were willing to pay significantly more for the overnight shipping of a chosen book. However, the relation was not found for all questions, perhaps due to how they were formulated, personal preferences affecting on the choices. For choices involving longer horizons, temporal preferences were weakly related or unrelated to CRT scores. Frederick's study detected as well, that the respondents scoring high perceived themselves to be significantly less impulsive, and, somehow paradoxically, less preoccupied with their future than the "low group". However, the "high group" perceived themselves to be more concerned about inflation. Choices between different monetary rewards are

problematic because a knowledgeable respondent can take into consideration for example interest rates, inflation and the ability to invest (Frederick, Loewenstein, & O'Donoghue, 2002).

Thinking dispositions assess people's propensities or tendencies that can facilitate reflective judgments (Basile & Toplak, 2015). They can be viewed as cognitive styles (Stanovich et al., 2016). The Consideration of Future Consequences Scale (CFC; Strathman, Gleicher, Boninger, & Edwards, 1994) is a thinking disposition measure, that has been associated with better rational thinking performance (Toplak et al., 2007; West, Toplak, & Stanovich, 2008) as well as with temporal discounting (Joireman, Balliet, Sprott, Spangenberg, & Schultz, 2008). The original Consideration of Future Consequences (CFC-12) is a 12-item scale measuring to what extent individuals consider distant outcomes when choosing their present behaviour. When the relation of CFC-12 and four different time preference measures were examined, the results showed that the CFC-12 was significantly associated with all four measures (Basile & Toplak, 2015). The relation of CFC-12 and CRT has been examined as well, and a small, but significant correlation detected (Toplak et al., 2011, 2014). The CFC-scale has been revised, and the CFC-14 scale applied in this study distinguishes between two subscales: concern with future consequences, CFC-Future, and concern with immediate consequences, CFC-Immediate (Joireman, Shaffer, Balliet, & Strathman, 2012). Presumably, the CFC-14 has not been previously used to detect the relations of CRT and time preference.

3.2. Risk preference

The relations of numeracy to other cognitive skills have been largely researched (Stanovich et al., 2016). However, the relation between risk preferences and cognitive ability had not been widely studied, which inspired Frederick (2005) to examine the relation of risk preference and the Cognitive Reflection Test (CRT) results. Some studies have reported a correlation between cognitive ability and risk preference (Beauchamp, Cesarini, & Johannesson, 2017; Benjamin et al., 2013; Burks, Jeffrey, Lorenz, & Aldo, 2009; Dohmen, Falk, Huffman, & Sunde, 2010), but some others did not detect a correlation (Brandstätter & Güth, 2002; Brañas-Garza, Guillen, & Lopez Del Paso, 2008). It is also questionable, whether the interpretation of risk attitude can be thought of as a personality trait because it is not stable across situations for most individuals (Bromiley & Curley, 1992). The same person can show different degrees of risk-taking in financial, career,

health, safety, ethical, recreational, and social decisions (Figner & Weber, 2011; Hanoch, Johnson, & Wilke, 2006; MacCrimmon & Wehrung, 1986).

According to prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1974), people are more willing to take risks to avoid losses, but more prone to avoid risks when there is a possibility to achieve sure gains. For example, when asked "*would you prefer 900 euros for sure OR 90% possibility to win 1000 euros*", most of the respondents would choose the sure gain. However, when asked whether you "*prefer to lose 900 euros for sure OR a 90% possibility to lose 1000 euros*", they would choose the risky option. The prospect theory predicts that people underweight outcomes that are merely probable, and overweight low probabilities. When people are winning, they become risk-averse, and when they are losing, they become risk-seeking.

A few attempts to study the relationship between risk preference and analytical thinking skills has been made. To assess the relation between CRT performance and risk preference, Frederick (2005) included 18 measures of risk preferences in his questionnaires. He documented that the respondents achieving a high CRT score acted opposite to prospect theory. Subjects achieving high scores were more prone to gamble for items involving gains, but less risk-seeking in the domain of losses. Oechssler and his colleagues (2009) replicated Frederick's study to some extent. They applied only two questions and divided the respondents into two groups including all the participants, whereas Frederick compared the results of the "extreme" achievers, the lowest achievers and the highest achievers. For both of the question items, Oechssler et al. found the high CRT group preferring slightly more the risky option in the domain of gains but being slightly more risk aversive in the domain of losses. With Finnish university students, the present study attempts to replicate the findings of original study of Frederick (2015).

4. Temperament as punishment and reward sensitivities

There have been unexpectedly few studies examining the influence of personality traits on decision-making biases. Research on the relationship between analytical thinking and temperament or personality traits is almost non-existent. This is surprising because, from the very beginning of developing the Cognitive Reflection Test (CRT), it has been assumed that success in the test requires an ability to suppress the intuitive answer, that is, individuals with high CRT scores would be more patient (Frederick, 2005). Also, a correlation between CRT and the Consideration of Future Consequences Scale (CFC-

12) has been detected (Toplak et al., 2011, 2014), suggesting, that the ability to analytical thinking is somehow related to farsightedness. Humans tend to act like cognitive misers, relying on Type 1 (heuristic) thinking only, and self-control is needed to induce further reflection. Self-control is known to be related to Conscientiousness, a personality trait that includes responsibility, industriousness, and orderliness (Moffitt et al., 2011).

In their study, Alós-Ferrer, Garagnani, and Hügelschäfer (2016), explored the possible influence of personality factors on both choices and response times in standard questions from the decision-making literature, including CRT. They applied standard personality scales such as the Big Five (McCrae & Costa, 1987). Their findings showed that the differences in response times were as predicted by dual-process theories: the CRT questions led to faster responses than other type of questions. However, even though personality trait Conscientiousness led to faster responses and Extraversion to slower ones, neither had a significant effect on the correctness of the responses. Agreeableness led to significantly more correct answers and Neuroticism and Openness to experience to more errors, but none of them had a significant effect on response times (Alós-Ferrer et al., 2016).

Temperament and personality can, of course, be studied from various perspectives. In this study, temperament is defined as a neurobiological basis of personality that accounts for inborn differences in individuals' typical ways of reacting to environmental stimuli (McNaughton & Corr, 2008; Rawlings, Tapola & Niemivirta, 2017). When temperament interacts with the environment, relatively stable emotional, motivational and behavioural patterns are formed (Rothbart, 2007). In this study, temperament is measured as punishment and reward sensitivity (Torrubia, César, Moltó, & Caseras, 2001; Rawlings et al, 2017), when the behavioural reaction to avoid aversive or approach appetitive stimuli is stemming from the innate behavioural inhibition (BIS) and behavioural approach system (BAS) (Corr, 2008; Elliot & Thrash, 2002; Corr & McNaughton, 2008). BIS is a self-protective function which detects signals of threatening punishments, whereas BAS is a function detecting signals for possible rewards (Carver & White, 1994; Corr, 2008; Corr & McNaughton, 2008; Elliot & Thrash, 2002). The BAS is often viewed as a more complex construction than the BIS (Corr, 2008), consisting of several dimensions (Carver & White, 1994; Colder et al., 2011).

Not only the sensitivity to punishment and reward, but also the sources of punishment and reward may vary from one individual to another (Colder et al., 2011; Rawlings et al., 2017). In their two sub-studies, Rawlings et al. (2017) observed a four-factor structure of

temperamental sensitivities (punishment and reward sensitivities) based on previous literature and previous empirical research. In their first study, they compiled a scale to cover sensitivity to punishment, two reward sensitivity scales reflecting differences in sources of intraindividual reward (novelty-seeking and positive expressiveness) and sensitivity to interindividual reward. The punishment sensitivity consists of sensitivity to failure, shyness, withdrawal and avoidance of difficult or novel situations. The interindividual reward sensitivity consists of sensitivity to external rewards, such as attention and praise, and seeking to impress others. The intraindividual reward sensitivity scales reflect two different sources of rewards: 1) enjoyment and seeking of novelty and 2) a tendency for excitement and open expression of positive emotions about personal successes. In their second study, the structure was the most part replicated.

Could analytical thinking skills, that is, the tendency to override the intuitive, incorrect answer and to engage in further reflection (Toplak et al., 2011), somehow reflect spontaneous processing of rewards and punishments? In order to better understand the mechanisms of individual differences in analytical thinking, the relation of analytical thinking and neurobiological temperament, assessed as punishment and reward sensitivities, is investigated in this study.

5. Research questions and assumptions

The main objective of this study was to look for relations between analytical thinking skills and selected psychological characteristics. Therefore the links between *analytical thinking skills* (assessed as performance in the Cognitive Reflection Test, CRT, Frederick, 2005) and individual differences in *time preference* (assessed as performance in the Consideration of Future Consequences Test, CFC; Joireman et al., 2012), *risk preference* (assessed as a questionnaire formerly used by Frederick, 2005) and *temperament* (assessed as sensitivity to punishment and rewards, Rawlings et al., 2017) were analysed.

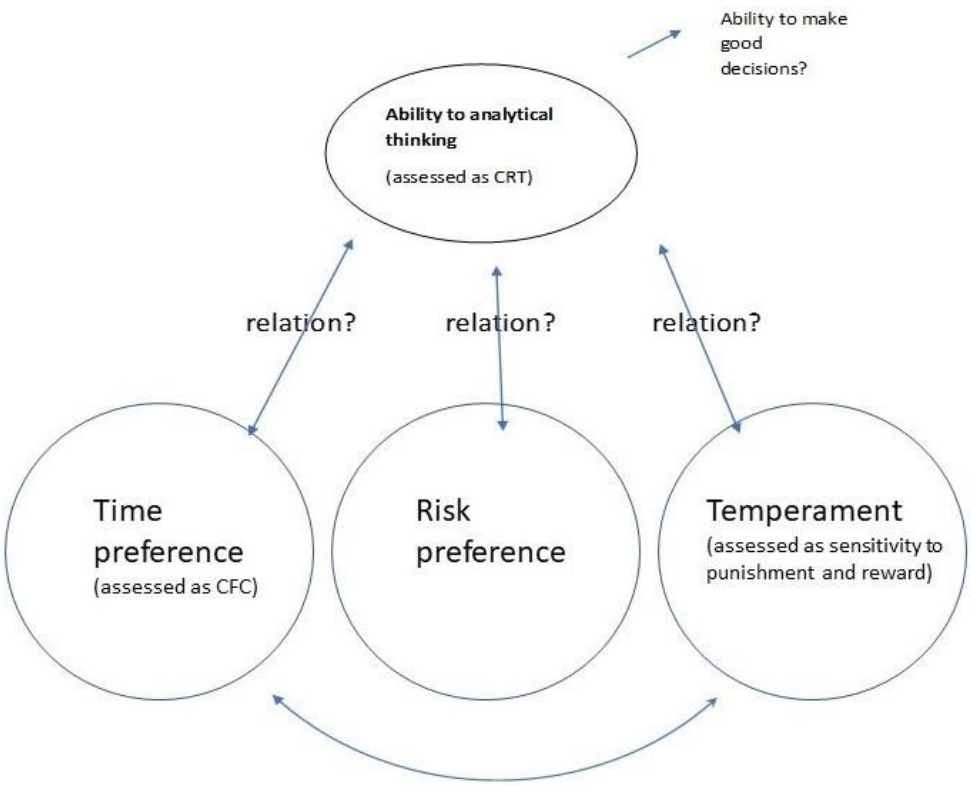

Figure 1. Research framework: Relations of CRT, time preference, risk preference and temperament.

This study investigated the following research questions:

a) Is there a relationship between the performances in the Cognitive Reflection Test (CRT) and in the time preference (CFC-14)? If a relationship exists, does CFC-14 score predict CRT score? Based on previous research it was expected, that individuals scoring high in CFC-Future should be more likely to gain a higher score in the CRT than those scoring low, and those scoring high on CFC-Immediate should be more likely to gain lower score in the CRT. In previous studies positive correlations have been detected between CFC-12 and CRT (Toplak et al., 2011, 2014). However, Toplak et al. (2014) reported, that CRT was not a very good predictor of CFC.

b) Is there a relationship between the performance in the CRT and temperament?

If a relationship exists, does temperament predict the CRT score? There has been little research examining the link between CRT and temperament or personality. A link between personality, assessed as the Big Five -personality scale (McCrae & Costa, 1987), and CRT has been examined (Alós-Ferrer et al., 2016), but the effects were not significant. The link between CRT and neurobiological temperament assessed as punishment and reward sensitivities (Rawlings et al., 2017) has not presumably been previously studied.

c) Is there a relationship between the performance in the CRT and risk preference? Based on previous research it was expected, that individuals scoring high in the CRT should prefer a risky option in the domain of gains but be more risk aversive in the domain of losses (Frederick, 2005; Oechssler et al., 2009).

d) How are the selected psychological characteristics, that is, time preference and temperament, related? Is there a link between temperament and the tendency to focus on the future or immediate consequences of one's behaviour?

6. Method

6.1. Participants and procedure

The sample used in this study consisted of 135 students (128 males, 7 females, M age = 21.55, SD = 1.52), practically representing the entire first-year class of National Defence University (NDU) in Finland. No background variables were available for the study. The participation was compulsory for the students. The participants were gathered in an auditorium where they completed a self-report questionnaire assessing analytical thinking skills, time preference, risk preference and temperament via a web link (Webprobol TM). The language used in the questionnaire was Finnish. The respondents were permitted to use a calculator or a mobile phone calculator if needed. They were informed that the gathered data will be treated confidentially and anonymously and that their identity cannot be determined.

The National Defence University provides higher education in the military field and trains officers for the Finnish Defence Forces. The students have completed their military service, and they are selected after a multi-tier selection process including evaluation of

achievement in prior education, achievement in military service, an exam phase containing a series of psychological and ability tests, and physiological screening. The annual intake of the NDU was approximately 140 at the time the data was gathered, and only 3–6 % of students were female, which is a typical gender distribution for this university.

6.2. Tasks and measures

To examine *the analytical thinking skills*, the Cognitive Reflection Test (CRT) was taken from Toplak, West and Stanovich (2014), and composed of seven problems (e.g. “A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? ____ cents”; see Appendix A).

The test for *time preference* was adapted from Joireman et al. (2012). The original Consideration of Future Consequences Test (CFC-12) is a unidimensional scale (Strathman et al., 1994). However, several studies have argued that a two-factor model, distinguishing between *CFC-Future* and *CFC-Immediate* subscales, explains the phenomena more efficiently (Joireman, Balliet, & Sprott, 2008; Petrocelli, 2003; Rappange, Brouwer, & van Exel, 2009; Toepoel, 2010). Therefore, a revised two-factor CFC-14 scale (Joireman et al., 2012) was applied in this study (see Appendix B). The CFC-14 subscales assess concern with future consequences (CFC-Future: 7 items, e.g. “Often I engage in a particular behavior in order to achieve outcomes that may not result for many years”) and concern with immediate consequences (CFC-Immediate: 7 items, e.g. “I only act to satisfy immediate concerns, figuring the future will take care of itself”).

The *risk preference* questionnaire (see Appendix C) was adapted from Frederick (2005) and consists of three categories of questions: 1. a certain gain versus a higher expected value (e.g. “1000 for sure or a 90% chance of 5000”); 2. a certain gain versus a lower expected value (e.g. “100 for sure or a 25% chance of 200”); 3. a certain loss versus a lower expected value (e.g. “lose 10 for sure or a 90% chance to lose 50”).

The scale for *temperament* was adapted from Rawlings, Tapola and Niemivirta (2017) and composed of four scales: a scale to cover *Sensitivity to punishment* (5 items; e.g. “I get upset easily if I am criticized or told off”), a scale to cover *Intraindividual reward sensitivity – positive expressiveness* (2 items; e.g. “I express my excitement and enjoyment openly, when I succeed at something”), a scale to cover *Intraindividual reward sensitivity – novelty-seeking* (3 items; e.g. “I will readily seek out novel situations”), and a scale to

cover external, *Interindividual reward sensitivity* (4 items; e.g. “I sometimes act hastily just to get an immediate reward or praise”; see Appendix D).

All scales were translated into Finnish. Items for CFC-14 and for temperament were mixed in the questionnaire. Responses to both scales were given on a seven-point Likert-type scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree).

6.3. Analyses

The statistical analyses were conducted by IBM SPSS version 23. The preliminary analyses consisted of examining the answers given in the Cognitive Reflection Test (CRT). An exploratory factor analysis was performed to test the construct validity of both of the models applied: Consideration of Future Consequences Test (CFC-14 model) and the temperament scales. Based on the results from the factor analyses, composite variables were constructed and Cronbach’s alpha values were calculated to evaluate the variables’ internal consistency and reliability (Field, 2015). The correlations between the three different scales were examined and standard regression analyses were performed to analyse the effect of the temperament on the CFC. Finally, the relationship between CRT scores and risk preference was examined with the cross-tabulation and Pearson chi-square tests.

7. Results

7.1. Preliminary analyses

7.1.1. CRT

Before proceeding to further analyses, the results from the CRT were analysed. The CRT consists of seven questions, CRT1-CRT7. The answers given in the CRT consist of three different categories: correct answer, intuitive (incorrect) answer and others (incorrect, but other than the intuitive) answer. Distribution of correct, intuitive and other answers are shown in table 1.

Table 1.

Distribution of answers per question.

(*N* = 135)

	correct %	intuitive %	other %
CRT1	55.6	37.8	6.7
CRT2	57.8	34.8	7.4
CRT3	71.9	21.5	6.7
CRT4	63.7	17.8	18.5
CRT5	45.2	43	11.9
CRT6	40.7	44.4	14.8
CRT7	82.2	11.1	6.7

Next, the distribution of CRT scores was calculated. The participants scored 0-7 points, receiving one point per correct answer. In this study, five participants (3.7%) scored 0 points, whereas 18 participants (13.3%) scored 7 points. Table 2 shows the distribution of CRT scores. Similarly, the participants were given one point per each intuitive answer, and an "intuitive score" (CRT_int) was calculated. As table 3 shows, only one participant gave the intuitive, incorrect answer for each of the seven CRT questions. 29 respondents (21.5%) gave none intuitive, incorrect answers. The descriptive statistics of the variables CRT and CRT_int were examined as well. Both variables were found to be distributed to an acceptable level, even though both exhibit some skewness (CRT: skewness -0.359, kurtosis -0.939; CRT_int: skewness 0.628, kurtosis -0.504), as shown in table 6.

Table 2.

Distribution of CRT scores (CRT). (*N* = 135)

Score	Frequency	Percent	Cumulative Percent
0	5	3.7	3.7
1	14	10.4	14.1
2	13	9.6	23.7
3	17	12.6	36.3
4	18	13.3	49.6
5	26	19.3	68.9
6	24	17.8	86.7
7	18	13.3	100.0
Total	135	100.0	

Table 3.

Distribution of “intuitive scores” (CRT_int).

(*N* = 135)

Score	Frequency	Percent	Cumulative Percent
0	29	21.5	21.5
1	34	25.2	46.7
2	21	15.6	62.2
3	20	14.8	77.0
4	14	10.4	87.4
5	11	8.1	95.6
6	5	3.7	99.3
7	1	0.7	100.0
Total	135	100.0	

7.1.2. Exploratory factor analyses

7.1.2.1. Consideration of Future Consequences

To examine the construct validity of the CFC-14 model (Joireman et al., 2012), an exploratory factor analysis was performed using Principal Axis Factoring with Direct Oblimin rotation. The following four items with communalities below 0.2 were removed*: “I generally ignore warnings about possible future problems because I think the problems will be resolved before they reach crisis level”* (communality 0.195); *“I think it is more important to perform a behaviour with important distant consequences than a behaviour with less-important immediate consequences”* (0.197); “*I think that sacrificing now is usually unnecessary since future outcomes can be dealt with at a later time”* (0.199); “*I only act to satisfy immediate concerns, figuring that I will take care of future problems that may occur at a later date”* (0.184).

The analysis showed two factors with eigenvalues greater than one (3.627, 1.697). The first factor corresponds to CFC-Future (CFC-F) (36.3% of the variance) and the second factor to CFC-Immediate (CFC-I) (17% of the variance), accounting for 61.5% of the total variance. The two-factor model was supported also by the scree test. There were no cross-loadings (see table 4).

Table 4.

Factor loadings of the CFC

	Factors	
	1	2
CFC-Future		
When I make a decision, I think about how it might affect me in the future.	0.80	
I consider how things might be in the future, and try to influence those things with my day to day behaviour.	0.67	
Often I engage in a particular behaviour in order to achieve outcomes that may not result for many years.	0.67	
My behaviour is generally influenced by future consequences.	0.62	
I think it is important to take warnings about negative outcomes seriously even if the negative outcome will not occur for many years.	0.51	
I am willing to sacrifice my immediate happiness or well-being in order to achieve future outcomes.	0.42	
CFC-Immediate		
I only act to satisfy immediate concerns, figuring the future will take care of itself.		0.75
Since my day to day work has specific outcomes, it is more important to me than behaviour that has distant outcomes.		0.62
My convenience is a big factor in the decisions I make or the actions I take.		0.56
My behaviour is only influenced by the immediate (i.e., a matter of days or weeks) outcomes of my actions.		0.41

7.1.2.2. Temperament

To examine the construct validity of the temperament scales (Rawlings et al., 2017), again an exploratory factor analysis was performed using Principal Axis Factoring with Direct Oblimin rotation. The following items with communalities below 0.2 were removed: *"I always try to get things or assignments done as quickly as possible"* (0.117) and "*I like activities or assignments where you can get positive feedback quickly and easily"* (0.103). As well item "*I often leave things undone just because I fear I will fail"* was removed due to consistency with previous studies.

The analysis showed four factors with eigenvalues greater than one (4.479, 2.145, 1.424, 1.015), as shown in table 5. The first factor corresponds to Sensitivity to punishment (32% of the variance), the second factor to Intraindividual reward sensitivity – positive expressiveness (15.3% of the variance), the third factor to Interindividual reward sensitivity (10.2% of the variance) and the fourth factor to Intraindividual reward sensitivity – novelty-seeking (7.3% of the variance), accounting for 64.7% of the total variance. One item, "*I will gladly be the centre of attention"*, cross-loaded onto three factors. It was included in the Interindividual reward sensitivity-factor, where it loaded with the highest value (0.39).

Table 5.

Factor loadings of the temperament scales

	Factors			
	1	2	3	4
Sensitivity to punishment SP				
I am easily shy in the company of people I don't know and in new situations.	-.97			
I feel very uncomfortable in new situations and places.	-.76			
I avoid talking or performing in public (e.g., at lectures).	-.59			
I withdraw easily in difficult or awkward situations.	-.57			
I get upset if I am criticised or told off.	-.38			
Intraindividual reward sensitivity – positive expressiveness SRiPE				
I express my excitement and enjoyment openly, when I succeed at something.		0.99		
I don't hold back my joy and enthusiasm when something nice happens to me.		0.70		
Interindividual reward sensitivity SRe				
I sometimes act hastily just to get an immediate reward or praise.			0.65	
I often do things just to be praised.			0.65	
I will gladly be the centre of attention.	0.31	0.32	0.39	
I often aim to impress other people.			0.35	
Intraindividual reward sensitivity – novelty-seeking SRiNS				
I will readily seek out novel situations.				0.78
I think it is exciting to get into new and surprising situations.				0.77
I get excited about new things easily.				0.53

7.2. Descriptive statistics and correlations

Based on the solutions from the factor analyses, composite variables were constructed. After creating composite scores, their normal distribution was examined and their internal consistency and reliability evaluated by calculating Cronbach's alpha values. The Cronbach's alphas were good/acceptable except for the factor Interindividual reward sensitivity (SRe). For SRe the alpha level was only .56, which must be taken into account when interpreting the results. The data was found to be suitable for parametric analyses. Descriptive statistics and Cronbach's alpha values are shown in table 6.

Table 6.

Descriptive statistics and Cronbach's alphas

Variable (*N* = 135)	*M*	*SD*	Alpha	Skewness	Kurtosis
CRT (7 items)	4.17	2.04	n/a	-0.359	-0.939
CRT_int (7 items)	2.1	1.79	n/a	0.628	-0.504
CFC-F (6 items)	5.24	0.80	0.79	-0.153	0.162
CFC-I (4 items)	2.94	0.91	0.69	0.266	-0.242
SP (5 items)	2.30	0.95	0.80	1.036	1.241
SRiPE (2 items)	5.03	1.12	0.82	-0.557	0.288
SRe (4 items)	3.29	0.79	0.56	0.026	-0.328
SRiNS (3 items)	5.33	0.94	0.77	-0.372	0.245

After constructing the composite scores for all variables shown in table 6, their correlations were examined with Pearson's correlation. CRT results did not show a considerable correlation with CFC or temperament. Negative correlations were found between Sensitivity to punishment (SP) and CFC-F (*r*= -.19), whereas Interindividual reward sensitivity SRe) and CFC-F (*r*= .22), and Intraindividual reward sensitivity – novelty-seeking (SRiNS) and CFC-F (*r*= .41) correlated positively. CFC-I correlated positively with Sensitivity to punishment (SP) (*r*= .41) and Interindividual reward sensitivity (SRe) (*r*= .27), and negatively with Intraindividual reward sensitivity – novelty-seeking (SRiNS) (*r*= -.20). All the correlations are shown in table 7.

Table 7.

Correlations between CRT scores, CFC and temperament

	CRT	CRT_int	CFC-F	CFC-I	SP	SRiPE	SRe	SRiNS
CRT	1							
CRT_int	-.89**	1						
CFC-F	.04	-.04	1					
CFC-I	-.05	.08	-.35**	1				
SP	.07	-.08	-.19*	.41**	1			
SRiPE	-.06	.15	.15	.05	-.19*	1		
SRe	.01	-.05	.22**	.27**	-.10	0.28**	1	
SRiNS	-.01	.05	.41**	-.20*	-.49**	0.39	0.22*	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

7.3. Temperament as a predictor of time preference

No correlations were found between the CRT performance and the other scales measured, but the temperament and time preference showed some significant correlations. Therefore, a multiple linear regression was calculated to predict the CFC-F and CFC-I from temperament. The results are shown in table 8.

For CFC-F, a significant regression equation was found (F (4,130) = 7.481, $p < 0.001$), with a R^2 of 0.187. Participant's predicted CFC-F score is equal to 3.010 – 0.029 (SRiPE) + 0.344 (SRiNS) + 0.153 (SRe) + 0.016 (SP), where only Intraindividual reward sensitivity – novelty-seeking (SRiNS) was a significant predictor.

For CFC-I, also a significant regression equation was found (F (4,130) = 12.421, $p < 0.001$), with a R^2 of 0.277. Participant's predicted CFC-I score is equal to 1.067 – 0.096 (SRiNS) + 0.360 (SRe) + 0.389 (SP) + 0.062 (SRiPE), where Interindividual reward sensitivity (SRe) and Sensitivity to punishment (SP) were significant predictors.

Table 8.

Regression of CFC on temperament

	SP			SRiNS			SRe			SRiPE		
	Beta	*t*	Sig.	Beta	*t*	Sig.	Beta	*t*	Sig.	Beta	*t*	Sig.
CFC-F	0.019	0.208	0.835	0.400	4.106	0.000	0.149	1.800	0.074	-0.041	-0.461	0.646
CFC-I	0.406	4.737	0.000	-0.099	-1.078	0.283	0.313	3.998	0.000	0.076	0.919	0.360

7.4. Risk preference analysis

Before analysing the relation between the CRT results and risk preference, the participants were grouped by their CRT scores. To simplify the analysis without affecting the conclusions, previous studies (Frederick 2005; Oechssler et al., 2009) have made a comparison between the lowest achieving group and the highest achieving group. Frederick (*N* = 3,428) applied a 3-item test and compared the results of the lowest performers (scoring 0 out of 3) with the highest performing group (scoring 3 out of 3). In this study (*N* = 135) a 7-item test was applied and therefore the groups were formed differently: "low group" consisted of 23.7% of the respondents, scoring 0, 1 or 2 out of 7. The "high group" consisted of 31.1%, scoring 6 or 7.

The relationship between CRT scores and risk preference was examined with the cross-tabulation and Pearson chi-square tests. Table 9 shows the risk-seeking behaviour among the two CRT groups (High/Low), the presentation being similar to Frederick's. The table shows the percentage of each group choosing the riskier option. For example, for the first item "*1000€ for sure or a 90 % chance of 5000€*", 71,9% from the "low group" and 81% from the "high group" chose the riskier option. For each item the higher percentage is highlighted in the table 9, as well as the probability values that are statistically significant or approaching significance.

Only a few of the results were significant. In the domain of certain gains versus higher expected value gambles, for two items the "low group" was more willing to gamble than the "high group", the other item being statistically significant ($p < 0.05$). For the rest of

the items the “high group” was more willing to gamble, only one of the results approaching significance. In the domain of certain gains versus lower expected value gambles, the “low group” was more willing to gamble for each item, but none of the results were significant.

For items involving losses, the “low group” was more willing to gamble for 4 out of 5 items. One of these results was statistically significant ($p < 0.05$). For one item the “high group” was more willing to gamble, and the result was statistically significant as well ($p < 0.05$).

Table 9.
Risk-seeking behaviour among low and high CRT groups

	Percentage choosing the riskier option	CRT group		
	Certain gains vs. Higher expected value gambles	Low	**High**	Stat.Signif.
1	1000€ for sure or a 90 % chance of 5000€	71.9%	**81%**	0.495
2	100€ for sure or a 90 % chance of 500€	**84.4%**	78.6%	**0.015**
3	1000€ for sure or a 75 % chance of 4000€	34.4%	**54.8%**	0.216
4	100€ for sure or a 75 % chance of 200€	**46.9%**	45.2%	0.919
5	100€ for sure or a 75 % chance of 150€	15.6%	**26.2%**	**0.087**
6	100€ for sure or a 50 % chance of 300€	12.5%	**28.6%**	0.249
7	500€ for sure or a 15 % chance of 1 000 000€	78.1%	**88.1%**	0.344
8	100€ for sure or a 3 % chance of 7 000€	9.4%	**21.4%**	0.309
	Certain gains vs. Lower expected value gambles	**Low**	High	Stat.Signif.
9	100€ for sure or a 25 % chance of 200€	**3.1%**	2.4%	0.964
10	100€ for sure or a 25 % chance of 300€	**6.3%**	4.8%	0.952
11	5€ for sure or a 4 % chance of 80€	**53.1%**	31%	0.157
12	5€ for sure or a 1 % chance of 80€	**25%**	14.3%	0.491
13	60€ for sure or a 1 % chance of 5000€	**21.9%**	14.3%	0.673

	Certain losses vs. Lower expected value gambles	**Low**	High	Stat.Sig-nif.
14	Lose 10€ for sure or a 90 % chance to lose 50€	9.4%	**16.7%**	**0.022**
15	Lose 100€ for sure or a 75 % chance to lose 200€	**28.1%**	16.7%	0.127
16	Lose 100€ for sure or a 50 % chance to lose 300€	**40.6%**	21.4%	0.202
17	Lose 50€ for sure or a 10 % chance to lose 800€	**56.3%**	42.9%	0.517
18	Lose 100€ for sure or a 3 % chance to lose 7000€	**68.8%**	42.9%	**0.046**

8. Discussion

The aim of the study was to look for links between analytical thinking skills, assessed as performance in the Cognitive Reflection Test (CRT), and some psychological characteristics that were assumed to correlate with it: time preference, risk preference, and temperament.

The preliminary results showed some variation in the difficulty of the CRT problems: 82.2% got CRT7 correct, and the second most easy question item was CRT3, which 71.9% got correct. CRT7 is differently formulated from the other test questions: the participant is asked to choose between three options (a, b and c), whereas all the other items are open-ended questions. CRT6 provoked the highest number of intuitive answers, 44.4%, the second being CRT5 with 43%. When the CRT scores were examined, 13.3% (*N* = 135) scored 7 points, and 17.8% scored 6 points. 21.5% did not choose the intuitive answer at all, and 25.2% chose it once.

In the original study of Frederick (2005), a large number of university students tested at Harvard, MIT and Princeton gave the intuitive answer for the bat-and-ball puzzle (CRT1) (Toplak et al., 2011). In another study, 85.6% of the participants from a large university gave the intuitive response for the CRT1 (Toplak et al, 2014). In this study, only 37.8% of the participants gave the intuitive answer for CRT1.

Also, in their earlier study (2011), Toplak and her colleagues reported 55.8% (*N* = 193) of the participants not solving any of the problems, and 6.6% solving all three items. In Frederick's study, 7% of the respondents from MIT scored 0, that is, did not solve any of the problems, and 48% scored 3, that is, solved all three problems. In the same study, 18% (*N* = 121) of the respondents from Princeton scored 0, and 26% scored 3. The

students from Harvard didn't do any better, 20% (*N* = 51) scored 0 whereas another 20% scored 3. In this study, only 3.7% (*N* = 135) scored 0 points. 13.3% of the Finnish participants solved all seven problems, and 17.8% solved six problems. Are the students of the Finnish National Defence University (NDU) less prone to act like cognitive misers than the students of the most prestigious universities in the U.S.?

There is a risk that the participants in this study were familiar with the original CRT problems (items CRT1-CRT3) or the objective of the test, because the test has attracted some publicity in the media during the years. It is also unknown how the selection processes of the universities differ, the selection process of the NDU being a multi-tier process including several different assessments. It is not known either, if the military training the students have received improves or encourages analytical thinking.

The gender distribution might offer an explanation for the difference in the results between the universities. The current data consisted of almost male participants only (128 males, 7 females). In his original study, Frederick (2005) showed that the test has a (male) gender bias: men score significantly higher than women, and women seem to make the intuitive mistake more often. Toplak et al. (2014) replicated Frederick's findings: there was a significantly better performance by males for both the original 3-item CRT as well as for the four new CRT-items introduced in their study.

8.1. Relations of the CRT with time preference and temperament

In previous studies, a relation between Cognitive Reflection Test performance and time preference has been documented (Frederick, 2015; Oechssler et al., 2009; Toplak et al., 2011, 2014). When proposed two alternatives, hypothetical choices between different amounts of money, the participants scoring higher in the CRT were found to be generally more "patient" (Frederick, 2005). Also, the same study revealed, that the high CRT group perceived themselves to be significantly less impulsive and more concerned about inflations than the "low group". However, ambiguously, the "high group" was less preoccupied with their future.

When the Consideration of Future Consequences scale was applied, CFC-12 and CRT showed a small, but significant correlation of 0.10 (Toplak et al., 2011). In their later study (2014), Toplak and her colleagues detected a correlation of 0.3 between CFC-12 and the new 7-item CRT, and a correlation of 0.21 between CFC-12 and the original 3-item

CRT. However, their regression analyses showed, that the 7-item CRT was not a very good predictor of time preference, explaining only 10% of the variance.

In this study, a revised CFC-14 scale (Joireman et al., 2012) was applied to measure the time preference. The CFC loaded onto two factors as expected, corresponding to CFC-Future (CFC-F) and CFC-Immediate (CFC-I). It was expected that individuals scoring high in CFC-Future should be more likely to gain a higher score in the CRT than those scoring low, and those scoring high on CFC-Immediate should be more likely to gain lower score in the CRT. The results did not support the research hypothesis. No correlation between CRT performance and CFC was found.

Also, the possible relation of CRT performance and temperament, assessed as punishment and reward sensitivities (Rawlings et al., 2017) was examined. The temperament loaded on four factors, corresponding to Sensitivity to punishment, Intraindividual reward sensitivity – positive expressiveness, Interindividual reward sensitivity and Intraindividual reward sensitivity – novelty-seeking. No correlation between CRT performance and temperament was found.

8.2. Relations of the CRT and risk preference

One of the research questions was: Are performance in the Cognitive Reflection Test and risk preference related? The prospect theory predicts, that people are more willing to take risks to avoid losses than to achieve gains, that is, more willing to gamble in the domain of losses and less willing to gamble in the domain of gains (Kahneman & Tversky, 1979; Tversky & Kahneman, 1974). In a previous study (Frederick, 2005) this was found to be true for the low CRT group, but not for the high CRT group.

Frederick's study detected, that the "high group" was more willing to gamble in the domain of gains, even when the gamble did not have a higher expected value and their choice, therefore, was less "rational". In this present study, in the domain of *higher expected value* gambles, the "high group" did show a stronger willingness to gamble in six out of eight items, but none of these results were significant, only one item approaching significance. Respectively for one item of the same category, the "low group" was more willing to gamble than the "high group", and the result was significant.

In the domain of certain gains versus *lower expected value* gambles none of the results were significant in this present study, nor in Frederick's study. In the current study, the "low group" showed a higher willingness to gamble for all the items, opposing Frederick's findings, where the "high group" was more willing to gamble for each item. In the domain of *certain losses versus lower expected value* gambles, the "low group" was more willing to gamble for 4 out of 5 items, corresponding to the results of the previous study (Frederick, 2005) where the "low group" was found to be more risk-seeking in the domain of losses. However, only one of the results was statistically significant, this being the item for which the "high group" was more willing to gamble than the "low group", making though a less rational choice.

Because of the lack of significant results, no conclusions for the risk preference of either group can be drawn from the results of this study, and the attempt to replicate results from earlier research failed. It is worth noticing, that the CRT groups were formed differently from prior research due to different version of CRT (3 items versus 7 items). A comparison only between the very lowest achievers (0 points) and the very highest achievers (7 points) would not have been fruitful due to low numerus ($N = 135$). Therefore the "low group" was formed from respondents scoring 0, 1 or 2 out of 7, and the "high group" from respondents scoring 6 or 7.

It is worth considering to what extent the correlation between cognitive ability and risk-taking is related to the ability to calculate the expected value. When monetary choices are offered, the respondent's knowledge about for example interest rates and inflation can have an affect on the choices (Frederick et al., 2002). The "high group" of this study was more prone to choose the option giving the higher expected value, that is, to make a more "rational" choice, in 15 of 18 items. In the original study of Frederick, the "high group" made the "rational" choice more often than the "low group", in 13 out of 18 items. The participants had the opportunity to use a calculator, which could have affected the results, even though the calculations of expected value are relatively easy. It is unknown, if calculators were available in the original studies. However, no conclusions can be drawn due to the lack of statistical significance. If only significant results and results approaching significance are taken into consideration, in the current study the "high group" was acting more "rationally" in two items and the "low group" made the more "rational" choice twice as well.

8.3. Relations of time preference and temperament

Several studies have examined relationships between the Consideration of Future Consequences (CFC-12; CFC-14) and health behaviour, for example, the habit of exercising (Adams & Nettle, 2009), limiting sun exposure (Heckman, Wilson, & Ingersoll, 2009) and the use of a condom (Appleby et al., 2005). A high level of CFC-Immediate seems to make an individual susceptible to self-control failure, whereas a high level of CFC-Future can buffer one against self-control failure (Joireman et al., 2008). People scoring high on the CFC-Future scale tend to adopt promotion orientation, whereas those scoring high on the CFC-Immediate tend to adopt a prevent orientation. People focusing on the future consequences should be more likely to adopt ideal self-goals (hopes, aspirations) and strive toward an abstract future goal. In contrast, people focusing on the immediate consequences of their actions should be more likely to adopt ought self-goals (duties and responsibilities), which also are more concrete and proximal. High CFC-Future predicts the tendency to seek larger, uncertain outcomes over smaller, certain outcomes (Joireman & Balliet, 2012) whereas high CFC-Immediate predicts the tendency to seek smaller, immediate outcomes over larger, delayed outcomes (Joireman et al., 2008).

One of the research questions was how are time preference and temperament related? Is there a link between temperament and the tendency to focus on the future or immediate consequences of one's behaviour? A weak, negative correlation between CFC-Future and Sensitivity to punishment was detected, whereas a moderate positive correlation between CFC-Immediate and Sensitivity to punishment (SP). Perhaps this indicates that an individual, who is trying to achieve a long-term goal, cannot be very concerned about how other people will see him or her in the forthcoming situations or "worry too much about the future" in general. The SP-scale consists of questions such as "*I feel very uncomfortable in new situations and places"* and "*I get upset if I am criticised or told off"*. Possibly a strong orientation in the future indicates that the individual is determined in aiming goals, more than having a distressed attitude towards the future.

CFC-Future correlated positively with Intraindividual reward sensitivity – novelty-seeking (SRiNS), where the correlation was found moderate, and with Interindividual reward sensitivity (SRe), where the correlation was weak. It should be kept in mind, that the reliability for the Interindividual reward sensitivity (SRe) was low (Cronbach's alpha of 0.56). Intra-

individual reward sensitivity – novelty seeking (SRiNS) was found to be a significant predictor of CFC-Future. It is, of course, easy to understand, that an individual who is keen to seek out novel situations and is excited about new things easily, has his or her thoughts in the future. Usually it is the future, that will bring new and exciting things, not so much your past.

CFC-Immediate correlated positively with Interindividual reward sensitivity (SRe) and negatively with Intraindividual reward sensitivity – novelty-seeking (SRiNS), both correlations were weak. One of items in SRe was *"I sometimes act hastily just to get an immediate reward or praise"*, and the correlation with novelty-seeking (SRiNS) was negative. Both findings are congruent with the attitude of valuing immediate satisfaction. Both Interindividual reward sensitivity (SRe) and Sensitivity to punishment (SP) were significant predictors of CFC-Immediate.

The findings were in line with previous research which states that people scoring high on the CFC-Future scale tend to adopt a promotion orientation, whereas those scoring high on the CFC-Immediate scale will tend to adopt a prevent orientation (Joireman et al., 2012).

8.4. Reliability and validity

The phenomena detected in previous studies were not replicated in this study and no relation between the CRT result and time preference or risk preference was found. This can be due to the relatively low numerus (N = 135), but also the differences in gender distribution, as discussed earlier in this chapter, need to be taken into consideration (128 males; 7 females).

The data was gathered from a university where the participation for the study was compulsory. The questionnaires were answered in Finnish, the mother tongue of all or most of the students. The participants identities were kept anonymous, and they cannot be identified when reading the present study. The research material has been handled securely, and the data will be stored in the National Defence University in Finland only. After publishing this study, all other data will be deleted.

The measurement scales CRT and CFC-14 have been used in numerous previous published studies. Different versions of the CFC-scale have been used in at least 125 publications (Joireman & King, 2016). It is also worth noting, that the respondents were students from a selective university. Therefore, the distribution of the success in the CRT, in particular, could have been different if the groups were formed from the general population.

8.5. Limitations and suggestions for further studies

Before closing, some limitations of the present study and suggestions for future studies will be considered. This study has several limitations. First, the number of participants could have been larger. Second, gender distribution could have been more equal. It would have been interesting to gather a sample from other universities in Finland, where the number of female students is higher. This would have enabled a better generalization of the results, a comparison between the sexes, and perhaps some relations would have been detected. Also, perhaps, this would have enabled a comparison between countries in the adult education context. However, for a masters' thesis, the data gathered was adequate.

Notwithstanding these limitations, the evidence reported in this study indicates that there is a relation between time preference, assessed as Consideration of Future Consequences (CFC-14) and neurological temperament, assessed as punishment and reward sensitivities. Future research with larger datasets could contribute to obtain a clearer picture of their possible links. Also, more research on the influence of temperament on decision-making and thinking styles is needed.

9. Conclusions

This study was not successful in replicating the results from previous studies, where a relation between the Cognitive Reflection Test (CRT) score and time preference, as well as risk preference, has been detected. Compared to many previous studies, in this study the data gathered was smaller, and gender distribution was different. Frederick (2005) found a considerable difference between men and women in CRT results. Men scored significantly higher than women on the CRT, even when SAT math scores were controlled. Women seemed to make the intuitive mistake more often, whereas the men made a wider variety of errors. In this study, it was not possible to make a comparison between

the sexes because only seven out of 135 participants were female. However, the gender distribution is a potential explanation for the reason why the students at the Finnish National Defence University seemed to be less prone to give intuitive, incorrect answers, than the students at prestigious universities in the U.S.

The CRT is no more the only available test of analytical thinking. At the end of the year 2016, Stanovich, West, and Toplak presented the first prototype for the Comprehensive Assessment of Rational Thinking (CART), which is designed to detect individual's Rationality Quotient (RQ). Unfortunately, their test was not yet available at the time the data for this research was gathered. The CART is a compilation of carefully selected subtests, for example, tests of probabilistic and statistical reasoning, argument evaluation, rational temporal discounting, sensitivity to expected value, risk knowledge, rejection of superstitious thinking and rejection of conspiracy beliefs (Stanovich et al., 2016). Therefore, it will still be tempting to continue to apply a much simpler test, such as the standard CRT.

One of the reasons for the interest of developing new tests for rational or analytical thinking is the ongoing discussion assuming that the psychometric value of the Cognitive Reflection Test necessarily declines with time. The CRT has become popular, the solutions are available online and even videos explaining how to solve the problems can be watched in YouTube. Perhaps the concerns have been exaggerated. Meyer and his colleagues examined 14,053 subjects who took the test up to 25 times (Meyer, Zhou, & Frederick, 2018). Their findings showed, that on average, respondents got only 0.024 additional items correct per exposure, and this small increase was driven entirely by the minority of subjects who continued to spend time reflecting on the items.

On top of this, the focus of studying and testing analytical thinking skills has been in detecting individual differences. There has been insufficient interest in the improvement of analytical thinking skills. Even if individual differences in the domain are related to psychological characteristics and are innate to some extent, the effort should be put on teaching and learning rational of thinking. Becoming aware of different heuristics and cognitive biases will certainly improve an individual's chances to make good decisions. It is believed, that people can become better thinkers (Evans & Frankish, 2009) and especially the skills to evaluate risks and probabilities should be thought at a young age (Gigerenzer, 2013). It would be interesting to apply these tests to evaluate the efficiency of teaching analytical thinking skills, especially in the adult education context among university students. This would not only benefit the students but could also lead to a better understanding of the mechanisms of analytical thinking.

References

Adams, J., & Nettle, D. (2009). Time perspective, personality and smoking, body mass, and physical activity: An empirical study. *British Journal of Health Psychology, 14*, 83–105. doi: 10.1348/135910708X299664

Alós-Ferrer, C., Garagnani, M., & Hügelschäfer, S. (2016). Cognitive Reflection, Decision *Biases, and Response Times. Frontiers in psychology, 7*, 1402. doi: 10.3389/fpsyg.2016.01402

Appleby, P.R., Marks, G., Ayala, A., Miller, L.C., Murphy, S., & Mansergh, G. (2005). Consideration of future consequences and anal intercourse among men who have sex with men. *Journal of Homosexuality, 50*(1), 119–133. doi: 10.1300/J082v50n01_06

Basile, A. G., & Toplak, M. E. (2015). Four converging measures of temporal discounting and their relationships with intelligence, executive functions, thinking dispositions, and behavioral outcomes. *Frontiers in Psychology, 6*, 728.

Beauchamp, J.P., Cesarini, D. & Johannesson, M.J. (2017). The Psychometric Properties of Measures of Economic Risk Preferences. *Journal of Risk and Uncertainty, 54*, 203–237. doi: 10.1007/s11166-017-9261-3

Bechara, A. (2004): The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition, 55*(1), 30–40. doi: 10.1016/j.bandc.2003.04.001

Benjamin, D.J., Brown, S.A., Shapiro, J.M. (2013). Who is 'Behavioral?' Cognitive Ability and Anomalous Preferences. *Journal of the European Economic Association, 11*(6), 1231–1255. doi: 10.1111/jeea.12055

Brandstätter, H., & Güth, W., (2002). Personality in Dictator and Ultimatum Games. *Central European Journal of Operations Research, 10*(3), 191–215.

Brañas-Garza, P., Guillen, P., & Lopez Del Paso, R. (2008). Math skills and risk attitudes. *Economics Letters 99*(2), 332–336. doi: 10.1016/j.econlet.2007.08.008

Bromiley, P., & Curley, S. P. (1992). Individual differences in risk taking. In J. F. Yates (Ed.), *Wiley series in human performance and cognition. Risk-taking behaviour*. Oxford, England: John Wiley & Sons. 87–132.

Burks, S. V., Jeffrey P. C, Lorenz G., & Aldo R. (2009). Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment. *Proceedings of the National Academy of Sciences, 106*(19), 7745–7750. doi: 10.1073/pnas.0812360106

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making, 5*(3), 182–191.

Carver, C.S. & White, T.L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology 67*(2), 319–333. doi: 10.13072/midss.536

Clore, G.L. & Huntsinger, J.R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399. doi: 10.1016/j.tics.2007.08.005

Colder, C.R, Trucco, E.M., Lopez, H.I., Hawk, Jr. L.W., Read, J.P., Lengua, L.J., Weiczorek, W.F. & Eiden, R.D. (2011). Revised reinforcement sensitivity theory and laboratory assessment of BIS and BAS in children. *Journal of Research in Personality 45* (2), 198–207.

Corr, P.J. (2008). Reinforcement Sensitivity Theory (RST): introduction. In Corr, P.J. (Ed.), *The Reinforcement Theory of Personality*. New York, NY, US: Cambridge University Press. 1–43.

Corr, P.J. & McNaugthon, N. (2008). Reinforcement sensitivity theory and personality. In Corr, P.J. (Ed.), *The Reinforcement Sensitivity Theory of Personality*. New York, NY, US: Cambridge University Press. 155–187.

Damasio, A. (1994). *Descartes' error*. London, England: Picador.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review, 100*(3), 1238–1260.

Dominowski, R. (1994). Insight and instructions. Presented at Annual Conference, British Psychological Society, Cognitive Psychology Section, 1-3. New Hall, Cambridge, UK.

Duckworth, A.L. (2011). The significance of self-control. *Proceeding of the National Academy of Sciences, 108*(7), 2639–2640. doi: 10.1073/pnas.1019725108

Elliot, A.J. & Thrash, T.M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology 82*(5), 804–818. doi: 10.1037/0022-3514.82.5.804

Evans, J. St. B.T., & Frankish, K. (Eds.) (2009). *In Two Minds*. *Dual Processes and Beyond*. Oxford, England: Oxford University Press.

Evans, J.S.B., & Stanovich, K.E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241. doi: 10.1177/1745691612460685

Fernbach, P.M., Sloman, S.A., Louis, R.S., & Shube, J.N. (2013). Explanation Fiends and Foes: How Mechanistic Detail Determines Understanding and Preference. *Journal of Consumer Research, 39*, 1115–1131. doi: 10.1086/667782

Field, A. (2015): *Discovering statistics using IBM SPSS Statistics* (4th ed.). London, England: SAGE Publication.

Figner, B., & Weber, E. U. (2011). Who Takes Risks When and Why?: Determinants of Risk Taking. *Current Directions in Psychological Science, 20*(4), 211–216. doi: 0.1177/0963721411415790

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*(4), 25–42. doi: 10.1257/089533005775196732

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature. 40*(2), 351–401. doi: 10.1257/002205102320161311

Gigerenzer, G. (2013): Smart Heuristics. 39-54. In Brockman, J. (Ed.): *Thinking. The New Science of Decision-Making, Problem-Solving, and Prediction.* New York, NY, US: HarperCollins Publishers.

Gilhooly, K.J., & Murphy, P. (2005). *Differentiating insight from non-insight problems. Thinking & Reasoning, 11*(3), 279–302. doi: 10.1080/13546780442000187

Hanoch, Y., Johnson, J. G., & Wilke, A. (2006). Domain Specificity in Experimental Measures and Participant Recruitment: An Application to Risk-Taking Behavior. *Psychological Science, 17*(4), 300–304. doi: 10.1111/j.1467-9280.2006.01702.x

Heckman, C. J., Wilson, D. B., & Ingersoll, K. S. (2009). The influence of appearance, health, and future orientations on tanning behavior. *American Journal of Health Behavior, 33*(3), 238–243. doi: 10.5993/AJHB.33.3.2

Joireman, J., Balliet, D., Sprott, D., Spangenberg, E., & Schultz, J. (2008). Consideration of future consequences, ego-depletion, and self-control: Support for distinguishing between CFC-Immediate and CFC-Future sub-scales. *Personality and Individual Differences, 45*(1), 15–21. doi: 10.1016/j.paid.2008.02.011

Joireman, J. & King, S. (2016). Individual Differences in the Consideration of Future and (More) Immediate Consequences: A Review and Directions for Future Research. *Social & Personality Psychology Compass, 10*(5), 313–326.

Joireman, J., Shaffer, M.J., Balliet, D., & Strathman, A. (2012). Promotion Orientation Explains Why Future-Oriented People Exercise and Eat Healthy: Evidence From the Two-Factor Consideration of Future Consequences-14 Scale. *Personality and Social Psychology Bulletin, 38*(10), 1272–1287. doi:10.1177/0146167212449362

Kahneman, D. (2012). *Thinking, Fast and Slow*. London, England: Penguin Books.

Kahneman, D., & Tversky, A. (1979): Prospect Theory: *An Analysis of Decision Under Risk. Econometrica, 47*(2), 263–91.

Kiss, H.J., I. Rodriguez-Lara, I., & Rosa-García, A. (2016). Think twice before running! Bank runs and cognitive abilities. *Journal of Behavioral and Experimental Economics*, *64*, 12–19.

Koehler, D.J., & James, G. (2010). *Probability Matching and Strategy Availability. Memory & Cognition, 38*(6), 667–676. doi: 10.3758/MC.38.6.667

LeDoux, J.E. (1996). The emotional brain: *The mysterious underpinnings of emotional life*. New York, US: Simon & Schuster.

Lerner, J.S., Li, Y., Piercarlo V. & Kassam, K.S. (2015): Emotion and Decision Making. *Annual Review of Psychology, 66*, 799–823. doi: 10.1146/annurev-psych-010213-115043

Liberali, J.M., Reyna, V.F., Furlan, S., Stein, L.M., & Pardo, S.T. (2012). Individual Differences in Numeracy and Cognitive Reflection, With Implications for Biases and Fallacies in Probability Judgment. *Journal of Behavioral Decision Making, 25*, 361–381. doi: 10.1002/bdm.752

Loewenstein, G. & Lerner, J.S. (2003). The role of affect in decision making. In Davidson R.J. et al. (Eds.): *Handbook of affective sciences*. New York, NY, US: Oxford University Press.

Mata, A., Ferreira, M. B., & Sherman, S. J. (2013). The metacognitive advantage of deliberative thinkers: A dual-process perspective on overconfidence. *Journal of Personality and Social Psychology, 105*(3), 353–373. doi: 10.1037/a0033640

Mazur, J.E. (1987). An adjusting procedure for studying delayed reinforcement. In Commons, V.M.L, Mazur, J.E., Nevin, J.A. & Rachlin, H. (Eds.), *Quantitative Analyses of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value*, Vol. V. 55–73. New York, NY, US: Psychology Press.

McClure, S.M., & Bickel, W.K. (2014). A dual-systems perspective on addiction: contributions from neuroimaging and cognitive training. *Annals of the New York Academy of Sciences, 1327*, 62–78. doi: 10.1111/nyas.12561

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90. doi: 10.1037/0022-3514.52.1.81

MacCrimmon K.R. & Wehrung D.A. (1986). Assessing Risk Propensity. In Daboni L., Montesano A., Lines M. (Eds.), *Recent Developments in the Foundations of Utility and Risk Theory.* Theory and Decision Library (An International Series in the Philosophy and Methodology of the Social and Behavioral Sciences), vol 47. Dordrecht, Netherlands: Springer. 291–309.

McNaughton, N. & Corr, P.J. (2008). The neuropsychology of fear and anxiety. In Corr, P.J. (Ed.), *The Reinforcement Sensitivity Theory of Personality*. New York, NY, US: Cambridge University Press. 44–94.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*(1), 3–19. doi: 10.1037/0033-295X.106.1.3

Meyer, A., Zhou, E. & Frederick, S. (2018). The non-effects of repeated exposure to the Cognitive Reflection Test. *Judgment and Decision Making, 13(3)*, 246–259.

Mischel, W. (1974). Processes in Delay of Gratification. In Berkowitz, L. (Ed.): *Advances in Experimental Social Psychology.* San Diego, California, US: Academic Press. 249–92.

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., Houts, R., Poulton, R., Roberts, B.W., Ross, S., Sears, M.R., Thomson, M.W, and Caspia, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceeding of the National Academy of Sciences 108 (7)*, 2693–2698.

Moritz, B.B., Hill, A.V., & Donohue, K. (2013). Individual Differences in the Newsvendor Problem: Behavior and Cognitive Reflection. *Journal of Operations Management, 31* (1-2), 72–85. doi: 10.1016/j.jom.2012.11.006

Nagin, S.S., & Pogarsky, G. (2003). An Experimental Investigation of Deterrence: Cheating, Self-Serving Bias, and Impulsivity. *Criminology, 41*(1), 501–27. doi: 10.1111/j.1745-9125.2003.tb00985.x

Narayanan, A., & Moritz, B.B. (2015). Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study. *Production and Operations Management Volume 24 (8)*, 1216–1234. doi: 10.1111/poms.12343

Obrecht, N.A., Chapman, G.B., & Gelman, R. (2009). An encounter frequency account of how experience affects likelihood estimation. *Memory & Cognition, 37*(5), 632–643. doi: 10.3758/MC.37.5.632

Oechssler, J., Roide, A., & Schmitz, P.W. (2009). Cognitive Abilities and Behavioral Biases. *Journal of Economic Behavior and Organization, 72*(1), 147–152. doi: 10.1016/j.jebo.2009.04.018

Parker, A.M., & Fischhoff, B. (2005). Decision-Making Competence: External Validation through an Individual-Differences Approach. *Journal of Behavioral Decision Making, 18* (1), 1–27. doi: 10.1002/bdm.481

Pennycook, G., Cheyne, J.A., Seli, P., Koehler, D.J., & Fugelsang, J.A. (2012). Analytic Cognitive Style Predicts Religious and Paranormal Belief. *Cognition, 123*(3), 335–346. doi: 10.1016/j.cognition.2012.03.003

Petrocelli, J. V. (2003). Factor Validation of the Consideration of Future Consequences Scale: Evidence for a Short Version. *The Journal of Social Psychology, 143*(4), 405–413. doi: 10.1080/00224540309598453

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability of delay. *Journal of the Experimental Analysis of Behavior, 55*(2), 233–244. doi: 10.1901/jeab.1991.55-233

Rappange, D. R., Brouwer, W. B. F., & Van Exel, N. J. A. (2009). Back to the Consideration of Future Consequences Scale: Time to reconsider? *The Journal of Social Psychology, 149*(5), 562–584. doi: 10.1080/00224540903232324

Rawlings, A. M., Tapola, A. & Niemivirta, M. (2017). Predictive effects of temperament on motivation. *International Journal of Educational Psychology, 6*, 148–182. doi: 10.17583/ijep.2017.2414

Rothbart, M.K. (2007). Temperament, development, and personality. *Current directions in Psychological Science, 16*(40), 207–212.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science, 300*(5626), 1755–8. doi: 10.1126/science.1082976

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., & Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *The Journal of Pediatrics, 162*(1), 90–93. doi: 10.1016/j.jpeds.2012.06.049

Shenhav, A., Rand, D.G., & Greene, J.D. (2012). Divine Intuition: Cognitive Style Influences Belief in God. *Journal of Experimental Psychology: General, 141*(3), 423–428. doi: 10.1037/a0025391

Shoda, Y., Mischel, W., & Peake, P.K. (1990). Predicting Adolescent Cognitive and Self-Regulatory Competencies form Preschool Delay of Gratification: Identifying Diagnostic Conditions. *Developmental Psychology, 26*(6), 978–86.

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Stanovich, K.E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT, US: Yale University Press.

Stanovich, K.E. (2011). *Rationality & the Reflective Mind*. New York, US: Oxford University Press.

Stanovich, K.E., & West, R.F. (2000). Individual Differences in Reasoning: Implications for the Rationality Debate? *Behavioral and Brain Sciences, 22*(5), 645–726.

Stanovich, K.E., West, R.F. and Toplak, M.E. (2016): *The rationality quotient. Toward a test of rational thinking*. Cambridge, MA, US: The MIT Press.

Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66*(4), 742–752. doi: 10.1037/0022-3514.66.4.742

Toepoel, V. (2010). Is consideration of future consequences a changeable construct? *Personality and Individual Differences, 48*(8), 951–956. doi: 10.1016/j.paid.2010.02.029

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K.E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making, 20*, 103–124. doi: 10.1002/bdm.544

Toplak, M.E., West, R.F., & Stanovich, K.E. (2011). The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-biases tasks. *Memory and Cognition, 39*(7), 1275–1289. doi: 10.3758/s13421-011-0104-1

Toplak, M.E., West, R.F., & Stanovich, K.E. (2014). Assessing Miserly Information Processing: An Expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147–168. doi: 10.1080/13546783.2013.844729

Torrubia, R., César, Á., Moltó, J. & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences, 31*(6), 837–862. doi: 10.1016/S0191-8869(00)00183-5

Tversky, A., & Kahneman, D. (1974): Judgment under Uncertainty: Heuristics and Biases. *Science, 185*, 1124–1131.

West, R.F., Toplak, M.E., & Stanovich, K.E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology, 100*(4), 930–941. doi: 0.1037/a0012842

Appendix

Appendix A
Cognitive Reflection Test (CRT)
Frederick (2005):

(1) A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? ____ cents [Correct answer 5 cents; intuitive answer 10 cents]

(2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ____ minutes [Correct answer 5 minutes; intuitive answer 100 minutes]

(3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ____ days [Correct answer 47 days; intuitive answer 24 days]

Toplak, West & Stanovich (2014):

(4) If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? _____ days [correct answer 4 days; intuitive answer 9]

(5) Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? _____ students [correct answer 29 students; intuitive answer 30]

(6) A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much has he made? _____ dollars [correct answer $20; intuitive answer $10]

(7) Simon decided to invest $8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon has: a. broken even in the stock market, b. is ahead of where he began, c. has lost money [correct answer c, because the value at this point is $7,000; intuitive response b].

Appendix B
Consideration of Future Consequences (CFC) Questionnaire
Joireman et al. (2012):

CFC-Future

1. When I make a decision, I think about how it might affect me in the future.

2. I consider how things might be in the future, and try to influence those things with my day to day behavior.
3. Often I engage in a particular behavior in order to achieve outcomes that may not result for many years.
4. I am willing to sacrifice my immediate happiness or well-being in order to achieve future outcomes.
5. I think it is important to take warnings about negative outcomes seriously even if the negative outcome will not occur for many years.
6. I think it is more important to perform a behavior with important distant consequences than a behavior with less important immediate consequences.
7. My behaviour is generally influenced by future consequences.

CFC-Immediate

8. I only act to satisfy immediate concerns, figuring the future will take care of itself.
9. My behavior is only influenced by the immediate (i.e., a matter of days or weeks) outcomes of my actions.
10. My convenience is a big factor in the decisions I make or the actions I take.
11. I generally ignore warnings about possible future problems because I think the problems will be resolved before they reach crisis level.
12. I think that sacrificing now is usually unnecessary since future outcomes can be dealt with at a later time.
13. I only act to satisfy immediate concerns, figuring that I will take care of future problems that may occur at a later date.
14. Since my day to day work has specific outcomes, it is more important to me than behavior that has distant outcomes.

Appendix C
Frederick (2005):
Certain gains vs. Higher expected value gambles
1. 1000 euros for sure or a 90 % chance of 5000 euros
2. 100 euros for sure or a 90% chance of 500 euros
3. 1000 euros for sure or a 75% chance of 4000 euros
4. 100 euros for sure or a 75% chance of 200 euros
5. 100 euros for sure or a 75% chance of 150 euros
6. 100 euros for sure or a 50% chance of 300 euros
7. 500 euros for sure or a 15% chance of 1 000 000 euros
8. 100 euros for sure or a 3% chance of 7000 euros

Certain gains vs. Lower expected value gambles
9. 100 euros for sure or a 25% chance of 200 euros
10. 100 euros for sure or a 25% chance of 300 euros
11. 5 euros for sure or a 4% chance of 80 euros
12. 5 euros for sure or a 1% chance of 80 euros
13. 60 euros for sure or a 1% chance of 5000 euros

Certain losses vs. Lower expected value gambles
14. Lose 10 euros for sure or a 90 % chance to lose 50 euros
15. Lose 100 euros for sure or a 75% chance to lose 200 euros
16. Lose 100 euros for sure or a 50% chance to lose 300 euros
17. Lose 50 euros for sure or a 10% chance to lose 800 euros

18. Lose 100 euros for sure or a 3% chance to lose 7000 euros

Appendix D
Rawlings, Tapola & Niemivirta (2017):
Temperament as reward and punishment sensitivities

Sensitivity to punishment
1. I am easily shy in the company of people I don't know and in new situations.
2. I feel very uncomfortable in new situations and places.
3. I avoid talking or performing in public (e.g., at lectures).
4. I withdraw easily in difficult or awkward situations.
5. I get upset easily if I am criticized or told off.

Intraindividual reward sensitivity – positive expressiveness
1. I express my excitement and enjoyment openly, when I succeed at something.
2. I don't hold back my joy and enthusiasm when something nice happens to me.

Interindividual reward sensitivity
1. I sometimes act hastily just to get an immediate reward or praise.
2. I often do things just to be praised.
3. I will gladly be the centre of attention.
4. I often aim to impress other people.

Intraindividual reward sensitivity – novelty-seeking
1. I will readily seek out novel situations.
2. I think it is exciting to get into new and surprising situations.
3. I get excited about new things easily.